

MINUTES OF THE
Task Force on Learning Standards and Accountability in Public Education
August 28, 2000 -8:30 a.m. - Room 223 State Capitol

Members Present:

Sen. Howard A Stephenson, Chair
Rep. Jeff Alexander, Chair
Sen. Karen Hale
Sen. L. Steven Poulton
Rep. Loraine Pace
Rep. Karen W. Morgan
Rep. LaWanna "Lou" Shurtliff
Rep. John Swallow
Jill Kennedy
Kim Burningham
Linda B. Ogden
Ila Rose Fife

Members Absent:

Lt. Gov. Olene S. Walker

Staff Present:

Mr. Bryant R. Howe, Research Analyst
Mr. O. William Asplund, Assistant Director
Mr. James Wilson, Associate General Counsel
Ms. Wendy Bangerter, Legislative Secretary

Note: A list of others present and a copy of materials distributed in the meeting are on file in the Office of Legislative Research and General Counsel.

1. Task Force Business –

MOTION: Rep. Swallow moved to approve the minutes of August 14, 2000. The motion passed unanimously.

2. Review of the Performance Standard Setting Process and Possible Implementation of the "Value Added" Assessment System in Utah – Dr. Barbara Lawrence, Coordinator, Testing and Evaluation, Utah State Office of Education, made a presentation on setting performance standards for the Utah Core Curriculum. She noted that the core curriculum now has content standards. These content standards define what students should know and the skills they should have at different points in their schooling. The various Criterion Referenced Tests attempt to determine the extent to which students have mastered those content standards. The next step is to establish performance standards for the CRTs and to set "cut scores" for the various proficiency levels, by test.

Performance standards facilitate the interpretation and use of assessment scores. Dr. Lawrence explained that in establishing performance standards, her office convened several groups of teachers to assist with the process. These teachers are familiar with student performance and the core curriculum at the various grade levels. These groups chose "mastery" as the basis for proficiency with the following proficiency levels: (1) below partial mastery; (2) partial mastery; (3) near mastery; and (4) mastery.

Dr. Lawrence explained that in determining the "cut scores" for each test that her office combined two well accepted methods: the "bookmark" method and "contrasting group study"

method. The performance standards-setting process is dependent upon the specific content of the curriculum so there may be different cut scores from grade level to grade level in the same subject area. Each time a new test is developed, new performance standards must be set for that test.

With the “bookmark” method, the test items are arranged in the test booklet in order of difficulty. “Readers” are then assembled who describe the desired performance of a student on curriculum subjects. A consensus is reached regarding how many items in the booklet a student needs to get correct in order to be considered to be at a given proficiency level. The percentage of items represented by placement of the bookmarks become the points where proficiency levels are divided. The “cut score” is the bottom point for a given proficiency level.

The “contrasting group study” occurs in the spring, prior to testing. Teachers are asked to judge each student’s anticipated level of mastery before the test is taken. After the student takes the test, a range of scores is assembled according to the teachers’ classification of student mastery, near mastery and partial mastery. Those scores are then compared with the cut scores generated by the bookmark process.

Dr. Lawrence reminded the task force members that the performance standard setting process was introduced to them last fall. She stated that they engaged in the process with the existing Criterion Referenced Test (CRT) series because the information was needed to comply with the requirements of HB 177 and the more times the process is completed, the more experienced the teachers are in making good judgements regarding cut scores. She emphasized that standard setting is very labor intensive. Once the cut scores are established, performance level descriptors are written to say what a student at a specific proficiency level can do in the language of the core curriculum in a particular subject area.

Ms. Lawrence related standards to accountability. She said that standards make it easier to monitor a student’s achievement from year to year or to act as a warning of missing skills in a student’s understanding. It helps teachers become “owners” of performance levels that students should be achieving. It creates a common goal throughout the state because all students will be measured against the same proficiency level.

Setting performance standards can also serve as a foundation for establishing acceptable levels of performance for schools. She noted that there is not a cut score level defined for those students who achieve above mastery because the core curriculum is written as a mastery curriculum. That does not limit schools, districts or teachers, however, from going beyond the core to challenge those students. She stated that she would be reporting the cut scores to the task force on the Oct 9th task force meeting for elementary language arts and elementary and secondary math and science.

Dr. Lawrence also reviewed the “Value-added” assessment model developed by Dr. William Sanders. This model uses assessments with specific characteristics and measures the growth of students over time, including a growth curve for an individual student. It allows students’ performance with a particular teacher to be considered, as well as the teacher’s performance over time. She reviewed the preferred pattern of student growth and indicated how it showed which teachers were meeting the needs of their students’ academic progress over the school year.

In his presentation to the local superintendents earlier this month, Dr. Sanders indicated that there are three criterion that must be met for his “value added” method to be applied to a particular test. First, the test must be related to the state’s core curriculum. Second, the test must be internally reliable. Third, the test must have sufficient “stretch.” That is, scores from various grade levels must be computed on the same scale.

Dr. Lawrence told the task force that while the Utah CRTs do relate to the core curriculum and are statistically reliable, they do not measure the materials on a common scale from year to year as Dr. Sanders requires. Rather, each CRT is matched to the core curriculum and the standards are set to each curriculum subject. She stated that if Utah were to adopt this model, the entire core curriculum and core assessment programs would have to be redesigned.

Wendy Bromley, President of Jordan School District Education Association, commented that the teachers who helped Dr. Lawrence’s official performance standard setting process were glad to be involved. She expressed concern with how some of the questions on the Algebra II CRT were written. She also expressed concern that the concepts being questioned on the test, according to some teachers she polled, were not top priority concepts to be taught. The Utah math core curriculum is very broad and not deep enough. Teachers have a difficult time covering all of the material specified by the state core curriculum.

Dr. Lawrence responded that the CRT is matched to the core curriculum for several reasons, one being to determine if the concepts in the core are being taught. This provides diagnostic information for teachers. She also stated that the content of the core was determined and written under a very systematic process. Ms. Bromley expressed concern that test results are not reported to teachers in a timely manner.

Bonnie Morgan, Coordinator, Core Curriculum, State Office of Education, agreed that the current math core curriculum has some difficulty with its wide breadth and shallow depth. She stated that the core is being revised to have five content standards instead of the previous 14, and is being aligned from grade level to grade level, considering what should be taught in preparation for the future grades. Classroom teachers are also heavily involved in developing the core curriculum. Textbooks are a guide line for teachers and they have the option to choose to teach

whatever concepts are needed by each group that comes into their classroom, they do not have to teach each chapter from start to finish. She stated that once the CRT results are returned to a school district, it is its responsibility to distribute the results to individual teachers so they can be utilized for lesson planning.

Dr. Lawrence noted that the addition of constructed response questions to the CRTs will increase the time required to report test results. The goal of the assessment system is to help teachers know what parts of the core curriculum that students have and have not mastered. The current performance standard setting process is being undertaken with the current core curriculum in order to satisfy the requirements of HB 177 and to get themselves and the teachers familiar with the process until such time as the new curriculum is developed. It was emphasized that more teacher development is needed to improve teacher skills and ability to teach the content of the core curriculum.

3. Review of Accountability Models Used in Other States -- Mr. Bryant R. Howe, Research Analyst, summarized the progress of the task force over the past year. He noted that the task force adopted the content standards as they currently found in the Utah Core Curriculum. The performance standards will be reported to the task force at its October 9, 2000 meeting.

Mr. Howe briefly reviewed the school ranking systems used in Florida, Texas, and North Carolina. Each state uses a different approach and different category labels.

4. Task Force Discussion on Assessments and Performance – Rep. Jeff Alexander lead a discussion regarding the task force goals to: (1) help students master content standards of the Utah Core Curriculum; and (2) help educators to be more effective.

Linda Ogden asked if this will be another unfunded mandate from the state. Rep Alexander stated that it is simply an exercise to identify schools with exemplary performance and schools who require additional assistance to help more students master the content standards..

Jill Kennedy said that the task force needs to examine the current assessment tools and the indicators of student behavior listed in HB 177 and determine which ones will be used..

Sen. Stephenson stated that the report card for the elementary school ought to include the data elements listed in HB 177. The question for the task force is how to apply some type of designation on whether the school has done a good job.

Rep. Morgan questioned the need for a designation label. She said that simply reporting the data will tell parents and constituents what they really want to know: how many students are

mastering the core curriculum. The report should indicate both the number and percent of students at each level.

Ms. Kennedy noted that with the Stanford 9 test you could show the school's actual percentile ranking compared to its expected percentile ranking. The task force should not use the labels of mastery, etc. as labels for schools. While these labels are appropriate for an individual assessment tool, they are not appropriate as labels for schools or districts.

Rep. Alexander argued for designations of schools. He stated that we give grades to students and that we should also give them to schools.

Rep. Morgan said that the task force should consider if applying designations are really needed to reach the goals and to provide the necessary interventions. She asked if labeling were really necessary to motivate a school to change.

Sen. Stephenson stated that whether it is necessary to apply designations to schools is the crux issue that the task force is considering this year. He distributed information on a new school designation law in Colorado and a report on the results of the Texas accountability system. In Florida, some of the schools that received a grade of "F" were very motivated to improve. He emphasized that Utah school districts have been reporting Stanford 9 data for 11 years. However, reporting just the raw data has not been as motivating as it could have been, if some type of designation system was established.

Superintendent Darlene Robles, Salt Lake City School Districts, said that labeling the schools will not change the way she responds to test scores. However, it will take more time to respond to media inquiries about school rankings.

Ila Rose Fife noted that the state's demographics and mobility are changing. You cannot necessarily compare current with past performance, given all the changes in our population. Schools have been doing comparatively well in the past 10 years, given the changes in the students that schools are serving. When a school obtains low scores, you do see teachers and administrators responding.

Rep. Swallow said that he hasn't yet made a final judgment on how to best identify a school. It is important to remember that a designation is simply a way to reflect the current status of the students in a given school – it is not a permanent label. The task force should weigh both the possible positive and negative consequences of a designation system. While the task force should not adopt an unfair system, it should be willing to make the tough choices that are necessary to improve our schools. He stated that we shouldn't be overly concerned with public reaction.

Linda Ogden emphasized that teachers want to be accountable for factors that they can control. The task force needs to provide help for every school because every school will have students who need help. We should identify students, not schools.

Rep. Pace said that the task force's goal is to help students master the content standards. The CRTs will tell us the extent to which students are mastering the content standards of the core curriculum. Maybe some other indicators should be added. With regards to other states, Utah has some unique characteristics that make it so some other states are not applicable to Utah. As long as we are getting the right information to the right individuals, then we can talk about motivation.

MOTION: Sen. Stephenson moved that a sample report card, including the data already being required by HB 177, be brought to the next meeting for task force consideration and to aid in discussion.

Linda Ogden said that labeling schools can lead to undue pressure from the community who does not understand what is going on. It also leads to dropouts and cheating.

Rep. Shurtliff said that the reporting of all the data is more useful than using a letter grade – both for the public and for the teachers.

The motion passed unanimously with Senators Hale and Poulton absent for the vote.

5. Other Business --

a. Next Meeting: Monday, September 11, 2000

MOTION: Rep. Morgan moved to adjourn the meeting. The motion passed unanimously. The meeting was adjourned at 11:56 a.m.